



1 AAAAAGAAAG GAAGAAAATG GAAATACAAC AAACACACCG CAAAATCAAT  
51 CGCCCTCTGG TTTCTCTCGC TTTAGTAGGA GCATTAGTCA GCATCACACC  
101 GCAACAAAGT CATGCCGCCT TTTTCACAAC CGTGATCATT CCAGCCATTG  
151 TTGGGGGTAT CGCTACAGGC ACCGCTGTAG GAACGGTCTC AGGGCTTCTT  
201 AGCTGGGGGC TCAAACAAGC CGAAGAAGCC AATAAAACCC CAGATAAACC  
251 CGATAAAGTT TGGCGCATTC AAGCAGGAAA AGGCTTTAAT GAATTCCTTA  
301 ACAAGGAATA CGACTTATAC AGATCCCTTT TATCCAGTAA GATTGATGGA  
351 GGTGTTGGGATT GGGGGAATGC CGCTAGGCAT TATTGGGTCA AAGGCGGGCA  
401 ACAGAATAAG CTTGAAGTGG ATATGAAAGA CGCTGTAGGG ACTTATACCT  
451 TATCAGGGCT TAGAACTTT ACTGGTGGGG ATTTAGATGT CAATATGCAA  
501 AAAGCCACTT TACGCTTGGG CCAATTCAAT GGCAATTCTT TTACAAGCTA  
551 TAAGGATAGT GCTGATCGCA CCACGAGAGT GATTTCAACG CTAAAAATAT  
601 CTCAATTGAT AATTTTGCGA AAATCAACAA CTCGTGTGGG TTCTGGAGCC  
651 GGGAGGAAAG CCAGCTCTAC GGTTTTGACT TTGCAAGCTT CAGAAGGGAT  
701 CACTAGCGAT AAAAACGCTG AAATTTCTCT TTATGATGGT GCCACGCTCA  
751 ATTTGGCTTC AAGCAGCGTT AAATTAATGG GTAATGTGTG GATGGGCCGT  
801 TTGCAATACG TGGGAGCGTA TTTGGCCCCT TCATACAGCA CGATAAACAC  
851 TTCAAAAGTA ACAGGGGAAG TGAATTTTAA CCACCTCACT GTTGGCGATA  
901 AAAACGCCGC TCAAGCGGGC ATTATCGCTA ATAAAAAGAC TAATATTGGC  
951 AACTGGATT TGTGGCAAAG CGCCGGGTTA AACATTATCG CTCCTCCAGA  
1001 AGGTGGCTAT AAGGATAAAC CCAATAATAC CCCTTCTCAA AGTGGTGCTA  
1051 AAAACGACAA AAATGAAAGC GCTAAAAACG ACAAACAAGA GAGCAGTCAA  
1101 AATAATAGTA AACTCAGGT CATTAAACCA CCCAATAGTG CGCAAAAAAC  
1151 AGAAGTTCAA CCCACGCAAG TCATTGATGG GCCTTTTGCG GGC GGCAAAG  
1201 ACACGGTTGT CAATATCAAC CGCATCAACA CTAACGCTGA TGGCAGGATT  
1251 AGAGTGGGAG GGTTTAAAGC TTCTCTTACC ACCAATGCGG CTCATTGCA  
1301 TATCGGCAAA GGCGGTGTCA ATCTGTCCAA TCAAGCGAGC GGGCGCTCTC

**FIG. 1A**



1351 TTATAGTGGA AAATCTAACT GGAATATCA CCGTTGATGG GCCTTTAAGA  
1401 GTGAATAATC AAGTGGGTGG CTATGCTTTG GCAGGATCAA GCGCGAATTT  
1451 TGAGTTTAAG GCTGGTACGG ATACCAAAAA CGGCACAGCC ACTTTTAATA  
1501 ACGATATTAG TCTGGGAAGA TTTGTGAATT TAAAGGTGGA TGCTCATACA  
1551 GCTAATTTTA AAGGTATTGA TACGGGTAAT GGTGGTTTCA ACACCTTAGA  
1601 TTTTAGTGGC GTTACAGACA AAGTCAATAT CAACAAGCTC ATTACGGCTT  
1651 CCACTAATGT GGCCGTAAA AACTTCAACA TTAATGAATT GATTGTAAA  
1701 ACCAATGGGA TAAGTGTGGG GGAATATACT CATTTTAGCG AAGATATAGG  
1751 CAGTCAATCG CGCATCAATA CCGTGCGTTT GGAAACTGGC ACTAGGTCAC  
1801 TTTTCTCTGG GGGTGTTAAA TTAAAGGTG GCGAAAAAT GGTATAGAT  
1851 GAGTTTTACT ATAGCCCTTG GAATTATTTT GACGCTAGAA ATATTAAAA  
1901 TGTTGAAATC ACCAATAAAC TTGCTTTTGG ACCTCAAGGA AGTCCTTGGG  
1951 GCACATCAAA ACTTATGTTT AATAATCTAA CCCTAGGTCA AAATGCGGTC  
2001 ATGGATTATA GCCAATTTTT AAATTTAACC ATTCAAGGGG ATTTCAATCA  
2051 CAATCAAGGC ACTATCAACT ATCTGGTCCG AGGTGGGAAA GTGGCAACCT  
2101 TAAGCGTAGG CAATGCAGCA GCTATGATGT TTAATAATGA TATAGACAGC  
2151 GCGACCGGAT TTTACAAACC GCTCATCAAG ATTAACAGCG CTCAAGATCT  
2201 CATTAAAAAT ACAGAACATG TTTTATTGAA AGCGAAAATC ATTGGTTATG  
2251 GTAATGTTTC TACAGGTACC AATGGCATTG GTAATGTAA TCTAGAAGAG  
2301 CAATTCAAAG AGCGCCTAGC CCTTTATAAC AACAATAACC GCATGGATAC  
2351 TTGTGTGGTG CGAAATACTG ATGACATTAA AGCATGCGGT ATGGCTATCG  
2401 GCGATCAAAG CATGGTGAAC AACCTGACA ATTACAAGTA TCTTATCGGT  
2451 AAGGCATGGA AAAATATAGG GATCAGCAAA ACAGCTAATG GCTCTAAAAT  
2501 TTCGGTGTAT TATTTAGGCA ATTCTACGCC TACTGAGAAT GGTGGCAATA  
2551 CCACAAATTT ACCCACAAC AGCACTAGCA ATGCACGTTC TGCCAACAAC  
2601 GCCCTTGCAC AAAACGCTCC TTTCGCTCAA CCTAGTGCTA CTCCTAATTT  
2651 AGTCGCTATC AATCAGCATG ATTTTGGCAC TATTGAAAGC GTGTTTGAAT

**FIG. 1B**



3/14

2701 TGGCTAACCG CTCTAAAGAT ATTGACACGC TTTATGCTAA CTCAGGCGCT  
2751 CAAGGCAGGG ATCTCTTACA AACCTTATTG ATTGATAGCC ATGATGCGGG  
2801 TTATGCCAGA AAAATGATTG ATGCTACAAG CGCTAATGAA ATCACCAAGC  
2851 AATTGAATAC GGCCACTACC ACTTTAAACA ACATAGCCAG TTTAGAGCAT  
2901 AAAACCAGCG GCTTACAAAC TTTGAGCTTG AGTAATGCGA TGATTTTAAA  
2951 TTCTCGTTTA GTCAATCTCT CCAGGAGACA CACCAACCAT ATTGACTCGT  
3001 TCGCCAAACG CTTACAAGCT TTAAGAGACC AAAAATTCGC TTCTTTAGAA  
3051 AGCGCGGCAG AAGTGTTGTA TCAATTTGCC CCTAAATATG AAAAACCTAC  
3101 CAATGTTTGG GCTAACGCTA TTGGGGGAAC GAGCTTGAAT AATGGCTCTA  
3151 ACGCTTCATT GTATGGCACA AGCGCGGGCG TAGACGCTTA CCTTAACGGG  
3201 CAAGTGGAAG CCATTGTGGG CGGTTTTGGA AGCTATGGTT ATAGCTCTTT  
3251 TAATAATCGT GCGAACTCCC TTAAGCTCTGG GGCCAATAAC ACTAATTTTG  
3301 GCGTGTATAG CCGTATTTTA ACCAACCAGC ATGAATTTGA CTTTGAAGCT  
3351 CAAGGGGCAC TAGGGAGCGA TCAATCAAGC TTGAATTTCA AAAGCGCTCT  
3401 ATTACAAGAT TTGAATCAAA GCTATCATTA CTTAGCCTAT AGCGCTGCAA  
3451 CAAGAGCGAG CTATGGTTAT GACTTCGCGT TTTTtagGAA CGCTTTAGTG  
3501 TTAAGACCAA GCGTGGGTGT GAGCTATAAC CATTTAGGTT CAACCAACTT  
3551 TAAAGCAAC AGCACCAATC AAGTGGCTTT GAAAAATGGC TCTAGCAGTC  
3601 AGCATTTATT CAACGCTAGC GCTAATGTGG AAGCGCGCTA TTATTATGGG  
3651 GACACTTCAT ACTTCTACAT GAATGCTGGA GTTTTACAAG AGTTCGCTCA  
3701 TGTTGGCTCT AATAACGCCG CGTCTTTAAA CACCTTTAAA GTGAATGCCG  
3751 CTCGCAACCC TTAAATACC CATGCCAGAG TGATGATGGG TGGGGAATTA  
3801 AAATTAGCTA AAGAAGTGTT TTTGAATTTG GGC GTTGT TT ATTGACAA  
3851 TTTGATTTCC AATATAGGCC ATTCGCTTC CAATTTAGGA ATGAGGTATA  
3901 GTTCTAAAT ACCGCTCTTA AACCCATGCT CAAAGCATGG GTTTGAAATC  
3951 TTACAAAACA

FIG. 1C



4/14

1 MEIQQTHRKI NRPLVSLALV GALVSITPQQ SHAAFFTTVI IPAIVGGIAT  
51 GTAVGTVSGL LSWGLKQAE E ANKTPDKPDK VWRIQAGKGF NEFPNKEYDL  
101 YRSLSSKID GGWDWGNAAR HYWVKGGQON KLEVDMKDAV GTYTLSGLRN  
151 FTGGDLVDNM QKATLRLGQF NGNSFTSYKD SADRTTRVIS TLKISQLIIL  
201 QKSTTRVSG AGRKASSTVL TLQASEGITS DKNAEISLYD GATLNLASSS  
251 VKLMGNVWMG RLQYVGAYLA PSYSTINTSK VTGEVNFNHL TVGDKNAAQA  
301 GIIANKKTNI GTLDLWQSAG LNIIAPPEGG YKDKPNNTPS QSGAKNDKNE  
351 SAKNDKQESS QNNSNTQVIN PPNSAQKTEV OPTQVIDGPF AGGKDTVNI  
401 NRINTNADGT IRVGGFKASL TTNAHLHIG KGGVNLSNQA SGRSLIVENL  
451 TGNITVDGPL RVNNQVGGYA LAGSSANFEE KAGTDTKNGT ATFNNDISLG  
501 RFVNLKVDAAH TANFKGIDTG NGGFNTLDFS GVTDKVNINK LITASTNVAV  
551 KNFNINELIV KTNGISVGEY THFSEDIGSQ SRINTVRLET GTRSLFSGGV  
601 KFKGGEKLV I DEFYYSPWNY FDARNIKNVE ITNKLAFGPQ GSPWGTSKLM  
651 FNNLTGQNA VMDYSQFLNL TIQGDFINNO GTINYLVRRG KVALTSVGNA  
701 AAMMFNNDID SATGFYKPLI KINSAQDLIK NTEHVLLKAK IIGYGNVSTG  
751 TNGISNVNLE EQFKERLALY NNNNRMDTCV VRNTDDIKAC GMAIGDQSMV  
801 NNPDNYKYLI GKAWKNIGIS KTANGSKISV YYLGNSTPTE NGGNTTNLPT  
851 NTTSNARSAN NALAQNAPFA QPSATPNLVA INQHDFGTIE SVFELANRSK  
901 DIDTLYANS G AQGRDLLQTL LIDSHDAGYA RKMIDATSAN EITKQLNTAT  
951 TTLNNIASLE HKTSGLQTLS LSNAMILNSR LVNLSRRHTN HIDSFAKRLQ  
1001 ALKDQKFASL ESAAEVLYQF APKYEKPTNV WANAIGGTSL NNGSNASLYG  
1051 TSAGVDAYLN GQVEAIVGGF GSYGYSSFNN RANSLNSGAN NTNFGVYSRI  
1101 LTNQHEFD FE AQGALGSDQS SLNFKSALLQ DLNQSYHYLA YSAATRASYG  
1151 YDFAFFRNAL VLKPSVGVSY NHLGSTNFKS NSTNQVALKN GSSSQHLFNA  
1201 SANVEARYYY GDTSYFYMNA GVLQEFAHVG SNNAASLNTF KVNAARNPLN  
1251 THARVMMGGE LKLAKEVFLN LGVVYLHNL I SNIGHFASNL GMRYSF

FIG. 2





CTCCATTTTAAGCAACTCCATAGACCACTAAAGAACTTTTTTTGAGGCTATCTTTGAAA  
GCTTAATTATACATGCTATAGTAAGCATGACACACAAACCAAACTATTTTTAGAACGCTT  
TCAAAAAGATTCATTTCTTATTTCTTGTTCTTATTAAAGTTCTTTCATTTTAGCAAATTT  
CTTTTTTCAATATTAATAATGATTAATGAAAAAAAAAAAAATGCTTGATATTGTTGTAT  
TTGACACTAACAAGATACCGATAGGTATGAACTAGGTATAGTAAGGAGAAACAATGACT  
M T  
AATAATCTTCAAGTAGCTTTTTCTTAAAGTTGATAACGCTGTCGCTTCATACGATCCTGAT  
23 N N L Q V A F L K V D N A V A S Y D P D  
CAATTAAGGGAAGAATACTCCAATAAAGCGATCAAAAATCCTACCAAAAAGAATCAGTAT  
63 Q L R E E Y S N K A I K N P T K K N Q Y  
GAATCTTCCACAAAGAGCTTTCAGAAATTTGGGGATCAGCGTTACCGAATTTTCACAAGT  
103 E S S T K S F Q K F G D Q R Y R I F T S  
GAAAATATCATACAACCCCCTATCCTTGATGATAAAGAGAAAGCGGAGTTTTTGAAATCT  
143 E N I I Q P P I L D D K E K A E F L K S  
ATGGGCGTGTTTGATGAGTCCTTGAAAGAAAGGCAAGAAGCAGAAAAAAATGGAGAGCCT  
183 M G V F D E S L K E R Q E A E K N G E P  
GATGTCAAAGAAGCAATCAATCAAGAACCAGTTCCTCATGTCCAACCAGATATAGCCACT  
223 D V K E A I N Q E P V P H V Q P D I A T  
AATTTTTCTAAATTCACCTCTTGCGGATATGGAAATGTTAGATGTTGAGGGAGTCGCTGAC  
263 N F S K F T L G D M E M L D V E G V A D  
TTAATGGGGAGTCATAATGGCATAGAACCTGAAAAAGTTTCATTGTTGTATGGGGGCAAT  
303 L M G S H N G I E P E K V S L L Y G G N  
AACAATGTGGCTACAATAATTAATGTGCATATGAAAAACGGCAGTGGCTTAGTCATAGCA  
343 N N V A T I I N V H M K N G S G L V I A  
GGCTCACAACGAGCATTAAAGTCAAGAAGAGATCCAAAACAAAATAGATTTTCATGGAATTT  
383 G S Q R A L S Q E E I Q N K I D F M E F  
ACTGAGATTAAAGATTTCCAAAAAGACTCTAAGGCTTATTTAGACGCCCTAGGGAATGAT  
423 T E I K D F Q K D S K A Y L D A L G N D  
AATGGGGATTTGAGCTACACTCTCAAAGATTATGGGAAAAAAGCAGATAAGCTTTAGAT  
463 N G D L S Y T L K D Y G K K A D K A L D  
TATTCTAATTTCAAATACACCAACGCCTCCAAGAATCCCAATAAGGGTGTAGGCGTTACG

FIG. 4A



ATCTGTCCTATTGATTTGTTTTCCATTTTGTTCCTCATGTGGATCTTGTGGATCACAAAC 120  
CATGTGCTCACCTTGACTAACCATTTCTCCAACCATACTTTAGCGTTGCATTTGATTTCT 240  
TTGTTAATTGTGGGTAAAAATGTGAATCGTCCTAGCCTTTAGACGCCTGCAACGATCGGG 360  
AATGAGAATGTTCAAAGACATGAATTGACTACTCAAGCGTGTAGCGATTTTGTAGCAGTCT 480  
AACGAAACCATTGACCAACAACCACAAACCGAAGCGGCTTTTAACCCGCAGCAATTTATC 600  
N E T I D Q Q P Q T E A A F N P Q Q F I  
CAAAAACCAATCGTTGATAAGAACGATAGGGATAACAGGCAAGCTTTTGAAGGAATCTCG 720  
Q K P I V D K N D R D N R Q A F E G I S  
TTTTCAGACTTTATCAATAAGAGCAATGATTTAATCAACAAAGACAATCTCATTGATGTA 840  
F S D F I N K S N D L I N K D N L I D V  
TGGGTGTCCCATCAAAACGATCCGTCTAAATCAACACCCGATCGATCCGAAATTTTATG 960  
W V S H Q N D P S K I N T R S I R N F M  
GCCAAACAATCTTTTGCAGGAATCATTATAGGGAATCAAATCCGAACGGATCAAAAGTTC 1080  
A K Q S F A G I I I G N Q I R T D Q K F  
ACTGGTGGGGATTGGTTGGATATTTTTCTCTCATTATATTTGACAAAAACAATCTTCT 1200  
T G G D W L D I F L S F I F D K K Q S S  
ACCACCACCGACATACAAGGCTTACCGCCTGAAGCTAGAGATTTACTTGATGAAAGGGGT 1320  
T T T D I Q G L P P E A R D L L D E R G  
ATTGATCCCAATTACAAGTTCAATCAATTATTGATTCACAATAACGCTCTGTCTTCTGTG 1440  
I D P N Y K F N Q L L I H N N A L S S V  
GGTGGTCCTGGAGCTAGGCATGATTGGAACGCCACCGTTGGTTATAAAGACCAACAAGGC 1560  
G G P G A R H D W N A T V G Y K D Q Q G  
GGTGGTGAGAAAGGGATTAACAACCCTAGTTTTTATCTCTACAAGAAGACCAACTCACA 1680  
G G E K G I N N P S F Y L Y K E D Q L T  
CTTGACAAAATAATGCTAAATTAGACAACCTTGAGCGAGAAAGAGAAGGAAAAATTCCGA 1800  
L A Q N N A K L D N L S E K E K E K F R  
CGTATTGCTTTTGTTCATAAAAAAGACACAAACATTTCAGCTTTAATTACTGAGTTTGGT 1920  
R I A F V S K K D T K H S A L I T E F G  
AGGGAGAAAAATGTTACTCTTCAAGGTAGCCTAAACATGATGGCGTGATGTTTGTGAT 2040  
R E K N V T L Q G S L K H D G V M F V D  
AATGGCGTTTCCCATTTAGAGTAGGCTTTAACAAGGTAGCTATCTTTAATTTGCCTGAT 2160

FIG. 4B



503 Y S N F K Y T N A S K N P N K G V G V T  
TTAAATAATCTCGCTATCACTAGTTTCGTAAGGCGGAATTTAGAGGATAAACTAACCCT  
543 L N N L A I T S F V R R N L C D K L T T  
GAATTGGTTGGAAAACTTTAACTTCAATAAAGCTGTAGCTGACGCTAAAAACACAGGC  
583 E L V G K T L N F N K A V A D A K N T G  
CATTTAGAGAAAGAAGTAGAGAAAAAATTGGAGAGCAAAAGCGGCAACAAAAATAAAATG  
623 H L E K E V E K K L E S K S G N K N K M  
GCTAATAGAGACGCAAGAGCAATCGCTTACGCTCAGAATCTTAAAGGCATCAAAAGGGAA  
663 A N R D A R A I A Y A Q N L K G I K R E  
GAATTCAAAAATGGCAAAAATAAGGATTTTCAGCAAGGCAGAAGAACTAAAGCCCTT  
703 E F K N G K N K D F S K A E E T L K A L  
AATGCAGCTTTGAATGAATTCAAAAATGGCAAAAATAAGGATTTTCAGCAAGGTAACGCAA  
743 N A A L N E F K N G K N K D F S K V T Q  
AAAGTTGATAATCTCAATCAAGCGGTATCAGTGGCTAAAGCAACGGGTGATTTTCAGTAGG  
783 K V D N L N Q A V S V A K A T G D F S R  
CAAAAAATGAAAGTCTCAATGCTAGAAAAAATCTGAAATATATCAATCCGTAAAGAAT  
823 Q K N E S L N A R K K S E I Y Q S V K N  
AAAACTTTTCGGACATCAAGAAAGAGTTGAATGCAAACTTGGAAATTTCAATAACAAT  
863 K N F S D I K K E L N A K L G N F N N N  
CAAGCAGCTAGCCTTGAAGAACCCATTTACGCTCAAGTTGCTAAAAAGGTAAATGCAAAA  
903 Q A A S L E E P I Y A Q V A K K V N A K  
CCTTTGAAAAGGCATGATAAAGTTGATGATCTCAGTAAGGTAGGGCTTTCAAGGAATCAA  
943 P L K R H D K V D D L S K V G L S R N Q  
TTTGGCAATCTAGAGCAAACGATAGACAAGCTCAAAGATTCTACAAAACACAATCCCATG  
983 F G N L E Q T I D K L K D S T K H N P M  
TACGCTACTAACAGCCACATACGCATTAATAGCAATATCAAAAATGGAGCAATCAATGAA

FIG. 4C





N G V S H L E V G F N K V A I F N L P D  
AAAGGATTGTCCCCACAAGAAGCTAATAAGCTTATCAAAGATTTTTTGAGCAGCAACAAA 2280  
K G L S P Q E A N K L I K D F L S S N K  
AATTATGATGAAGTGAAAAAAGCTCAGAAAGATCTTGAAAAATCTCTAAGGAAACGAGAG 2400  
N Y D E V K K A Q K D L E K S L R K R E  
GAAGCAAAAGCTCAAGCTAACAGCCAAAAAGATGAGATTTTTGCGTTGATCAATAAAGAG 2520  
E A K A Q A N S Q K D E I F A L I N K E  
TTGTCTGATAAACTTGAAAATGTCAACAAGAATTTGAAAGACTTTGATAAATCTTTTGAT 2640  
L S D K L E N V N K N L K D F D K S F D  
AAAGGTTTCGGTGAAAGATTTAGGTATCAATCCAGAATGGATTTCAAAAGTTGAAAACCTT 2760  
K G S V K D L G I N P E W I S K V E N L  
GCAAAAAGCGACCTTGAAAATTCGGTTAAAGATGTGATCATCAATCAAAAGGTAACGGAT 2880  
A K S D L E N S V K D V I I N Q K V T D  
GTAGAGCAAGCGTTAGCCGATCTCAAAAATTTCTCAAAGGAGCAATTGGCCCAACAAGCT 3000  
V E Q A L A D L K N F S K E Q L A Q Q A  
GGTGTGAATGGAACCCTAGTCGGTAATGGGTTATCTCAAGCAGAAGCCACAACCTCTTTCT 3120  
G V N G T L V G N G L S Q A E A T T L S  
AACAATAATGGACTCAAAAACGAACCCATTTATGCTAAAGTTAATAAAAAGAAAGCAGGG 3240  
N N N G L K N E P I Y A K V N K K K A G  
ATTGACCGACTCAATCAAATAGCAAGTGGTTTGGGTGTTGTAGGGCAAGCAGCGGGCTTC 3360  
I D R L N Q I A S G L G V V G Q A A G F  
GAATTGGCTCAGAAAATTGACAATCTCAATCAAGCGGTATCAGAAGCTAAAGCAGGTTTT 3480  
E L A Q K I D N L N Q A V S E A K A G F  
AATCTATGGGTTGAAAGTGCAAAAAAAGTACCTGCTAGTTTGTGAGCGAAACTAGACAAT 3600  
N L W V E S A K K V P A S L S A K L D N  
AAAGCGACCGGCATGCTAACGCAAAAAAACCTGAGTGGCTCAAGCTCGTGAATGATAAG 3720

FIG. 4D



1023 Y A T N S H I R I N S N I K N G A I N  
ATAGTTGCGCATAATGTAGGAAGCGTTCCTTTGTCAGAGTATGATAAAATTGGCTTC  
1063 I V A H N V G S V P L S E Y D K I G F  
GTAAAAGACACTAATTCTGGCTTTACGCAATTTTAAACCAATGCATTTTCTACAGCA  
1103 V K D T N S G F T Q F L T N A F S T A  
GGTTTCCAAAAATCTTAAAGGATTAAGGAATACCAAAAACGCAAAAACCACCCCTTG  
1143 G F Q K S  
TGAATGCTACCAATTCATGGTATCATATCCCCATACATTTCGTATCTAGCGTAGGAAG  
AACTCTGTAAAATCCCTATTATAGGGACACAGAGTGAGAACCAACTCTCCCTACGG  
GACAGACACTAACGAAAGGCTTTGTTCTTTAAAGTCTGCATGGATATTTCCCTACCCC  
CGAAAATTAATTAAGGGTTATAAAGAGAGCATAACTAGAAAAACAAGTAGCTATA  
GAAAAATCAGAAAAACCATAGGAATTATCACACCTTATAATGCCCAAAAAAGACGCT  
ATGCCTTTCAAGGTGAAGAGGCAGATATTATTATTATTCACCGTGAAAACCTGTG  
ATCTCATTTTTGTGGGTAAAAAGTCTTTCTTTGAGAATTTATGAAGCGATGAGAAGA  
CATTCTTCGCTTCAAACGCTTTCATAAATCTCTCTAAAGCGCTTTATAATCAACAC  
TTATTAGCGTTACAATTTGAGCCATTCTTTAGCTTGTTTTCTAGCCAGATCACATC  
CTGCAAATATCCTACAATAGCATCGCCCGAATGGATGAGTAGGGGGGGTGTGAAAG  
TAAAATAATCACTTCGGGAAAATCTTTAAGGGAGTGAAATAATAACGCATGCAAGTT  
TGCGAAACATTCAAATAGCCTTGTTGTTTCAGGGCATTGTCATAAGCGTTGGATTGG  
GCTAAAATGCTTGGCTCAATCACGCCACAATAGGGATTTTGAATGCTTTTGCATC  
TTGAAAAAATCCAAAGCCTCTAAGCCAAATTGCTTGATCGTAGTGGGGTCTTTAGTG  
AGGCTTTTTAAACGCTAAACCCTCCACACCGCTATCAAAAACGCCTATTTTCATG  
TCTTCATTGTCCTTAGTTTGTTGCATTTTAGAATAGACAAAGCTT 5925

FIG. 4E



E K A T G M L T Q K N P E W L K L V N D K  
AACCAGAAGAATATGAAAGATTATTCTGATTCTGTTCAAGTTTTCCACCAAGTTGAACAATGCT 3840  
N Q K N M K D Y S D S F K F S T K L N N A  
TCTTATTACTGCTTGGCGAGAGAAAATGCGGAGCATGGAATCAAGAACGTTAATACAAAAGGT 3960  
S Y Y C L A R E N A E H G I K N V N T K G  
CTAAAGCGAGGGGTTTTTAACTCCTTAGCAGAAATCCCAATCGTCTTTAGTATTTGGGA 4080

TGTGCAAAGTTACGCCTTTGGAGATATGATGTGTGAGACCTGTAGGGAATGCGTTGGAGCTCA 4200  
GCAACATCAGCCTAGGAAGCCCAATCGTCTTTAGCGGTTGGGCACTTCACCTTAAAATATCCC 4320  
AAAAAGACTTAACCTTTGCTTAAATTAAGTTTGATTGTGCTAGTGGGTTGCTGCTATAGTG 4440  
ACAAAGATCAAGTTCAAAAAATCATAGAGCTTTTAGAGCAAATTGATCGCGCTCTTAACCAAA 4560  
TGCGATCAGAAGTGGA AAAATACGGCTTCAAGAATTTTGATGAGCTCAAATAGACACTGTGG 4680  
GTAATCTTTCTTTCTTGCTAGATTCTAAACGCTTGAATGTGGCTATTTCTAGGGCAAAAGAAA 4800  
ATATCTTTAGCGCTATTTTGCAAGTCTGTAGATAGGTAATCTTTTCAAAGATAATCATTAGA 4920  
AATACCCTTATAGTGTGAGCTATAGCCCCTTTTGGGAATTGAGTTATTTGACTTTAAATTT 5040  
GCCGCTCGCATGAAATTCACCTTTAGGGAATGCGTGTGCATTTTTTTTAAGGGCGTATTTTTG 5160  
GGCAAAATGCTCCATAAAATAGCCCTCAATTTTTTGAGCGATTAAGGGAAAATGCGTGCAACC 5280  
TCTAACAATTCGCCCTCTAAATACTTTCTTCAATCAAAGGCACAAAAGAGAAGTGGCTAAA 5400  
ATCGTCGCTTTTGTCCCTAGCACTAAAATAGGGGCGTTTTTATCTTTTACTTGTGCTTGATC 5520  
TCTTCTAAAGCTAGAGCGCTCGCTGTGTTGCATGCCACAATCAATAATTCAATCTGGTGCGGT 5640  
CCATAAGGCACTCTAGCCGTATCGCCATAATAGATGATTTTCATCAAATAATTGCGCTTTTAAA 5760  
ACACTTTTTTAATTTAATGGGATTAATTAGGGATTTTATTTTTTCATTCATTAAGTTTAAAAAT 5880

FIG. 4F



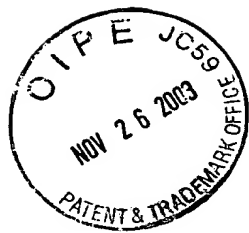
10 30 50  
AAGCTTGCTGTCATGATCACAAAAACACTAAAAACATTATTATTAAGGATACAAAATG  
M  
70 90 110  
GCAAAAGAAATCAAATTTTCAGATAGTGCGAGAAACCTTTTATTTGAAGGCGTGAGGCAA  
A K E I K F S D S A R N L L F E G V R Q  
130 150 170  
CTCCATGACGCTGTCAAAGTAACCATGGGGCCAAGAGGCAGGAATGTATTGATCCAAAAA  
L H D A V K V T M G P R G R N V L I Q K  
190 210 230  
AGCTATGGCGCTCCAAGCATCACCAAAGACGGCGTGAGCGTGGCTAAAGAGATTGAATTA  
S Y G A P S I T K D G V S V A K E I E L  
250 270 290  
AGTTGCCAGTAGCTAACATGGGCGCTCAACTCGTTAAAGAAGTAGCGAGCAAAACCGCT  
S C P V A N M G A Q L V K E V A S K T A  
310 330 350  
GATGCTGCCGGCGATGGCACGACCACAGCGACCGTGCTAGCTTATAGCATTTTAAAGAA  
D A A G D G T T T A T V L A Y S I F K E  
370 390 410  
GGTTTGAGGAATATCACGGCTGGGGCTAACCTATTGAAGTGAAACGAGGCATGGATAAA  
G L R N I T A G A N P I E V K R G M D K  
430 450 470  
GCTGCTGAAGCGATCATTAAATGAGCTTAAAAAGCGAGCAAAAAAGTAGGCGGTAAAGAA  
A A E A I I N E L K K A S K K V G G K E  
490 510 530  
GAAATCACCCAAGTGGCGACCATTCTGCAAACCTCCGATCACAATATCGGGAAACTCATC  
E I T Q V A T I S A N S D H N I G K L I  
550 570 590  
GCTGACGCTATGGAAAAAGTGGGTAAAGACGGCGTGATCACCGTTGAGGAAGCTAAGGGC  
A D A M E K V G K D G V I T V E E A K G  
610 630 650  
ATTGAAGATGAATTGGATGTCGTAGAAGGCATGCAATTTGATAGAGGCTACCTCTCCCCT  
I E D E L D V V E G M Q F D R G Y L S P

FIG. 5A



670 690 710  
TATTTTGTAAACGAACGCTGAGAAAATGACCGCTCAATTGGATAATGCTTACATCCTTTTA  
Y F V T N A E K M T A Q L D N A Y I L L  
730 750 770  
ACGGATAAAAAAATCTCTAGCATGAAAGACATTCTCCCGCTACTAGAAAAAACCATGAAA  
T D K K I S S M K D I L P L L E K T M K  
790 810 HindIII  
GAGGGCAAACCGCTTTTAATCATCGCTGAAGACATTGAGGGCGAAGCTTTAACGACTCTA  
E G K P L L I I A E D I E G E A L T T L  
850 870 890  
GTGGTGAATAAATTAAGAGGCGTGTTGAATATCGCAGCGGTTAAAGCTCCAGGCTTTGGG  
V V N K L R G V L N I A A V K A P G F G  
910 930 950  
GACAGAAGAAAAGAAATGCTCAAAGACATCGCTATTTTAACCGGCGGTCAAGTCATTAGC  
D R R K E M L K D I A I L T G G Q V I S  
970 990 1010  
GAAGAATTGGGCTTGAGTCTAGAAAACGCTGAAGTGGAGTTTTTAGGCAAAGCTGGAAGG  
E E L G L S L E N A E V E F L G K A G R  
1030 1050 1070  
ATTGTGATTGACAAAGACAACACCACGATCGTAGATGGCAAAGGCCATAGCGATGATGTT  
I V I D K D N T T I V D G K G H S D D V  
1090 1110 1130  
AAAGACAGAGTCGCGCAGATCAAAACCCAAATTGCAAGTACGACAAGCGATTATGACAAA  
K D R V A Q I K T Q I A S T T S D Y D K  
1150 1170 1190  
GAAAAATTGCAAGAAAGATTGGCTAAACTCTCTGGCGGTGTGGCTGTGATTAAAGTGGGC  
E K L Q E R L A K L S G G V A V I K V G  
1210 1230 1250  
GCTGCGAGTGAAGTGGAATGAAAGAGAAAAAAGACCGGGTGGATGACGCGTTGAGCGCG  
A A S E V E M K E K K D R V D D A L S A  
1270 1290 1310  
ACTAAAGCGGCGGTTGAAGAAGGCATTGTGATTGGTGGCGGTGCGGCTCTCATTGCGCGG  
T K A A V E E G I V I G G G A A L I R A

FIG. 5B



1330 1350 1370  
GCTCAAAAAGTGCATTTGAATTTGCACGATGATGAAAAAGTGGGCTATGAAATCATCATG  
A Q K V H L N L H D D E K V G Y E I I M  
1390 1410 1430  
CGCGCCATTAAAGCCCCATTAGCTCAAATCGCTATCAACGCTGGTTATGATGGCGGTGTG  
R A I K A P L A Q I A I N A G Y D G G V  
1450 1470 1490  
GTCGTGAATGAAGTAGAAAAACACGAAGGGCATTTTGGTTTTAACGCTAGCAATGGCAAG  
V V N E V E K H E G H F G F N A S N G K  
1510 1530 1550  
TATGTGGATATGTTTAAAGAAGGCATTATTGACCCCTTAAAAGTAGAAAGGATCGCTCTA  
Y V D M F K E G I I D P L K V E R I A L  
1570 1590 1610  
CAAAATGCGGTTTCGGTTTCAAGCCTGCTTTTAACCACAGAAGCCACCGTGCATGAAATC  
Q N A V S V S S L L L T T E A T V H E I  
1630 1650 1670  
AAAGAAGAAAAAGCGACTCCGGCAATGCCTGATATGGGTGGCATGGGCGGTATGGGAGGC  
K E E K A T P A M P D M G G M G G M G G  
1690 1710 1730  
ATGGGCGGCATGATGTAAGCCCGCTTGCTTTTTAGTATAATCTGCTTTTAAAATCCCTTC  
M G G M M •  
1750 1770 1790  
TCTAAATCCCCCCTTTCTAAAATCTCTTTTTTGGGGGGGTGCTTTGATAAAACCGCTCG  
  
1810 1830  
CTTGTA AAAACATGCAACAAAAAATCTCTGTTAAGCTT

FIG. 5C